Cuadrado Torre, A., et al. Bidirectional highway traffic for network simulation, in *2017 IEEE Vehicular Networking Conference (VNC), 27-29 Nov. 2017, Torino, Italy, pp. 77-80*

DOI: https://doi.org/10.1109/VNC.2017.8275628

Bidirectional Highway Traffic for Network Simulation

Alejandro Cuadrado Torre*, Marco Fiore†, Claudio Casetti*, Marco Gramaglia‡, Maria Calderón‡
*Politecnico di Torino, Torino, Italy
†CNR-IEIIT, Torino, Italy
‡University Carlos III of Madrid, Madrid, Spain

***Abstract*—Reliable and reproducible research in vehicular networking demands, among other requisites, a suitable representation of road traffic. We leverage real-world fine-grained measurement data recorded on the M40 highway in Madrid, Spain, to feed a realistic and properly parameterized microscopic simulation of vehicular mobility. The output is the first dataset of bidirectional highway traffic that is publicly accessible to the vehicular networking community. We leverage the dataset to demonstrate the validity in a complete highway scenario of the three-phase law of vehicular network connectivity, previously proven only on single carriageways.**

I. Introduction

Technologies enabling direct communications between vehicles will soon hit the market. Regulators are pushing dedicated short-range communication (DSRC) into the requirements for newly manufactured cars [1]. Also, the future generation of cellular networks is expected to support device-to-device communication (D2D), with automotive applications as a main use case [2]. In addition to services including road safety, road congestion prevention and management, liability attribution, localisation, or advertising, vehicle-to-vehicle communications are also anticipated to support improved automatic driving.

The evaluation of solutions based on vehicle-to-vehicle communication heavily relies on simulation, whose credibility rests, among other factors, on the realism of road traffic modeling. The availability of dedicated, reliable datasets of road traffic is thus paramount to the dependability and reproducibility of vehicular networking research.

Datasets of vehicular mobility that are publicly available and suitable for network simulation encompass urban areas [3]–[5] and highway environments [6], [7]. However, the latter are exclusively unidirectional, *i.e.*, only represent movements on a single carriageway. In this short paper, we introduce a first dataset of bidirectional highway traffic for network research.

We also showcase the utility of the dataset by studying how bidirectional traffic affects the connectivity of a network based on vehicle-to-vehicle communication. To this end, we consider the three-phase law of vehicular connectivity, which was previously demonstrated on single carriageways only [7], and prove that it holds also on bidirectional highways.

II. Measurement data: a glance at road traffic

In order to represent highway traffic in a realistic way, we leverage real-world measurement data collected on a highway segment. Our data were recorded on the M40, an orbital motorway of Madrid, Spain, by the Directorate General of Traffic (DGT), the Spanish national transportation agency. Magnetic induction loop pairs deployed at three different kilometric points on the M40 gathered timestamped information on the speed of each transiting vehicle. Different magnetic loops on each lane of both carriageways allowed monitoring vehicles on a per-lane basis, on both motorway directions.

Fig. 1a illustrates the exact locations where measurements were performed. They are named after their associated motorway kilometric points, **`51.2`**, **`52.2`** and **`57.0`**: there is thus 1 km between the first two points, and around 7 km separate the first and third point. Also, the carriageways heading North-East and South-West are indicated as **`C`** (for *creciente*, *i.e.*, increasing), and **`D`** (for *decreciente*, *i.e.*, decreasing), respectively; both have four lanes at all measurement points.

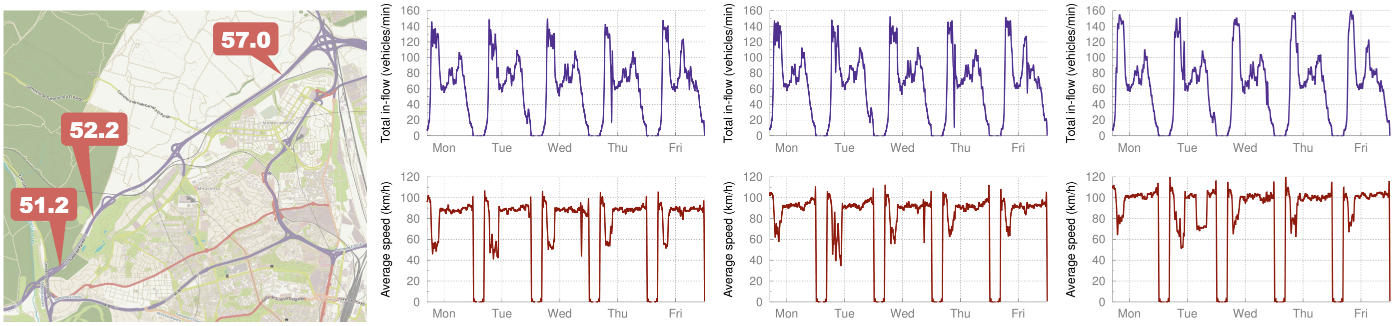
The data we use in our study were gathered during five consecutive working days, Monday to Friday, during a week in April 2015. On each day, measurements cover the period between 5 am and midnight, *i.e.*, negligible overnight traffic is not monitored. We stress that this is a much longer time period than those considered in previous works [6], [7]. The data are also very accurate: vehicle transit is recorded at every 100 ms, whereas typical road traffic monitoring systems aggregate statistics over minutes.

The top plot of Fig. 1b shows the in-flow at **`51.2 C`**, *i.e.*, the number of vehicles transiting by the first measurement point in the North-East direction, as captured during the five days covered by the data. The bottom plot displays instead the speed recorded at the same location and time. We observe a heavy traffic peak in the mornings of all days, which causes a significant reduction of speed. Denser traffic also characterises the hours around lunch and late in the afternoon, although without impact on the travel speed. These patterns reflect typical commuting behaviors on M40. Fig. 1c and Fig. 1d portray the same figures of merit, for **`52.2 C`** and **`57.0 C`**: the time series are nearly identical to those for **`51.2 C`**.

Interestingly, the trends change when considering the opposite travel direction. Fig. 2 shows the in-flow and speed observed at **`57.0 D`** during the same five days. Differences are remarkable: the traffic peak, although still apparent in the morning, has been shifted to the afternoon. As a result, road congestion and reduced speed appear later during the day. These observations capture the alternance of commuting, with most drivers travelling on the **`C`** direction in the morning and returing along the **`D`** direction in the afternoon.

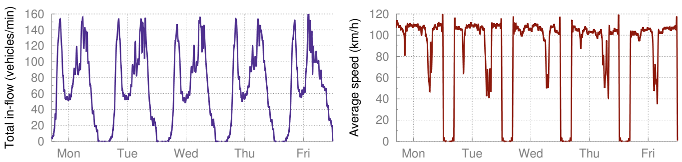
The previous results refer to statistics that are aggregate over all lanes. When considering each lane separately, we notice a substantial heterogeneity. For instance, Fig. 3 reports the temporal dynamics of in-flow and speed on separate lanes at **`51.2 C`**. Clearly, both measures change across lanes, with higher speed on the leftmost lane 4, higher in-flow on the center-left lane 3, and lowest speed on the rightmost lane 1. Again, these results reflect well real-world driving behaviors, such as increasing speeds from the rightmost to the leftmost lane or a tendency to stay in the central lane.

Overall, these result provide multiple examples of how our measurement data capture a variety of realistic phenomena in highway traffic, and are a sound basis for our investigation.



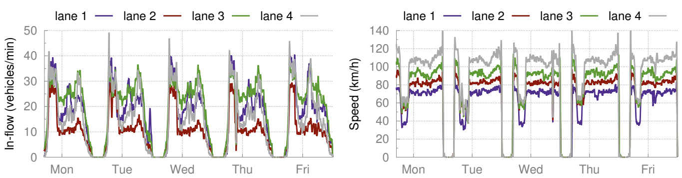
(a) Measurement collection locations (b) **51.2 C** (c) **52.2 C** (d) **57.0 C**

Fig. 1. (a) Geographical locations of the measurement points on the M40 motorway. (b,c,d) In-flow (in vehicles/minute, top) and speed (in km/h, bottom) recorded at the three subsequent measurement points **52.1 C**, **52.2 C**, and **57.0 C** during five days.



(a) **57.0 D**, in-flow (b) **57.0 D**, speed

Fig. 2. Time series of (a) in-flow and (b) speed recorded at measurement point **57.0 D** on the M40 motorway.



(a) **51.2 C**, in-flow (b) **51.2 C**, speed

Fig. 3. Time series of (a) in-flow and (b) speed recorded at measurement point **51.2 C** on the M40 motorway. Results are disaggregated by lane.

III. Mobility simulation: blending data & models

We feed the road traffic data to a simulator of microscopic vehicular mobility, so as to generate highway traffic datasets that describe the movement of each individual vehicle on the target motorway segment. This requires: *(i)* denoising the measurement data; *(ii)* verifying that a quasi-stationary state exists between the measurement points; and, *(iii)* parameterizing the simulation models so that they accommodate the real-world data. We discuss these three steps below.

A. Measurement data denoising

As typical of real-world data sources, our measurement data contain errors, mainly due to limitations of the magnetic induction loop technology. We identify and solve the following flaws:

Duplicated vehicles. Induction loops have a tendency to count twice the same vehicle, resulting in two transits at identical speed occurring in a very short timespan. Setting a one-second inter-arrival time threshold on subsequent vehicles with the same speed removes this error.

Inaccurate speed. Many transit events are associated with speeds that are plain unreasonable, *i.e.*, null or extremely high or low with respect to those of the preceding and following vehicles. We solve this problem by setting the speed of a vehicle to the average of those of the ten previous cars that transited on the same lane, whenever its original recorded speed exceeds twice that value or falls short of half of it.

TABLE I
Percentage of vehicles with erroneous data.

	Type	Mon	Tue	Wed	Thu	Fri
C	Duplicated vehicles	5.22%	5.92%	5.35%	5.32%	6.11%
	Inaccurate speed	2.28%	2.71%	1.92%	1.83%	2.38%
	Unsafe driving	0.30%	0.38%	0.34%	0.23%	0.30%
D	Duplicated vehicles	7.26%	6.57%	7.13%	6.88%	7.32%
	Inaccurate speed	3.18%	3.09%	2.80%	2.56%	2.97%
	Unsafe driving	0.18%	0.21%	0.15%	0.17%	0.16%

Unsafe driving conditions. These errors, which may partly be ascribed to the hazardous driving style of some people, result in vehicles travelling at very high speeds and very close to the vehicle in front. Independently of the source of the problem, no simulation model can handle such situations, which could cause accidents on real roads. We address them by fixing the velocity so that safe driving circumstances are restored, by transforming simple equations of linear dynamics. Specifically, we remove vehicles that travel exceedingly close to their front car, *i.e.*, which verify the following condition

$$v_{i-1} \cdot h_i < \ell,$$

where v_{i-1} is the speed of the front vehicle (in m/s) on the same lane, h_i is the headway time (in s) with respect to the same front vehicle, *i.e.*, the inter-arrival time between the current vehicle and the previous one on the same lane, and ℓ (set to 5 meters in our tests) is the typical length of a vehicle plus the minimum bumper-to-bumper distance. Also, we set the speed of a vehicle to that of the preceding car on the same lane if their combination of velocity and distance does not allow avoiding an accident at full braking deceleration, *i.e.*, if

$$v_{i-1} \cdot h_i - \frac{(v_i - v_{i-1})^2}{2b} < \ell,$$

where v_i is the speed of the target vehicle (in m/s), and b is the maximum deceleration (in m/s^2).

Tab. I summarizes the percentages of vehicles whose measurement data were fixed according to each of the actions presented above. Percentages remain reasonably low and are consistent across carriageways **C** and **D** and days, which warrants that real-word dynamics are preserved.

TABLE II
IN-FLOW AND SPEED CORRELATION BETWEEN SUBSEQUENT MEASUREMENT POINTS ON M40.

	Measurement points	Time bin length			
		10 s		30 s	
		In-flow	Speed	In-flow	Speed
C	51.2 $\rightarrow$ 52.2	0.836	0.862	0.924	0.950
	52.2 $\rightarrow$ 57.0	0.792	0.811	0.888	0.908
D	57.0 $\rightarrow$ 52.2	0.840	0.817	0.919	0.888
	52.2 $\rightarrow$ 51.2	0.850	0.812	0.932	0.895

B. *Quasi-stationarity of traffic on the considered M40 segment*

In Fig. 1–3, we observed how diversity is much higher across carriageways and lanes than when comparing different measurement points along the same direction. This lets us speculate that traffic conditions on the considered road segment may be quasi-stationary, *i.e.*, they are comparable between the in-flow (**51.2 C** and **57.0 D**, on each carriageway respectively) and out-flow (**57.0 C** and **51.2 D**, respectively). A very common assumption in vehicular network simulations, quasi-stationarity avoids that uncontrolled road traffic phenomena introduce a bias in the performance evaluation of the vehicular network.

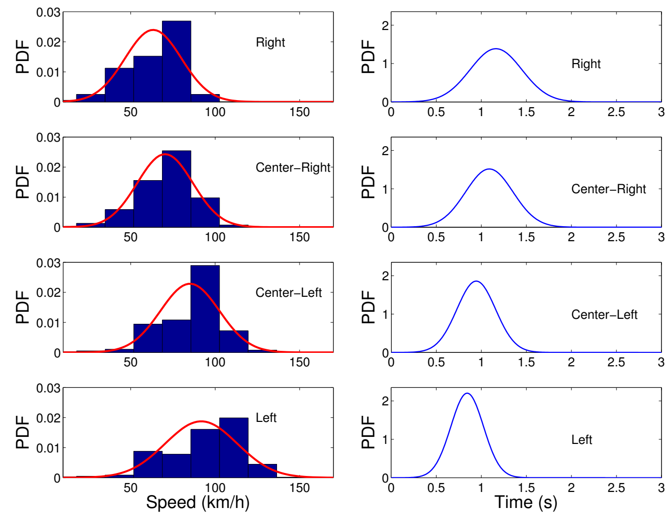
In order to verify if the road traffic conditions are stable in our scenario, we investigate the correlations between in-flow and speed recorded at the different measurement locations on the same direction. We thus bin the data recorded at each measurement point over 10 and 30 seconds time intervals, and compute the average in-flow and speed on each bin. We then observe how such values compare across subsequent measurement points, for the **C** and **D** directions separately. In doing so, we take into account the time needed to travel from one measurement location to the next one, using the current average speed: *e.g.*, if the average speed at **52.2 C** is 100 km/h, the in-flow and speed at such point will be compared with those computed at **57.0 C** 180 seconds (the time needed to travel the 5 km separating the two kilometric points) later.

Tab. II summarizes the Pearson correlation coefficients. Values are farily high in all cases, and range between 0.8 and 0.95 depending on the time bin length and measurement points pair. We conclude that quasi-stationarity is an acceptable approximation for the road traffic conditions on the considered M40 segment. Interestingly, this result implies that a couple of interchanges with in- and out-ramps between the measurement points do not affect significantly the observed traffic dynamics.

C. *Parameterizing microscopic mobility models*

Our simulator implements well-known and widely adopted models of individual drivers' behavior: the car-following Intelligent Driver Model (IDM) [8] and the lane-changing model MOBIL [9]. These microscopic-level models have a number of parameters: default values based on experiments are available, but they are the same for all vehicles and do not reflect the heterogeneous driving styles encountered in real-world measurement data.

In order to properly parameterize the models, we resort to the approach proposed in [7]. We thus maintain the default values for the maximum acceleration and deceleration, minimum vehicle distance, politeness factor when overtaking, acceleration biases favoring use of the rightmost free lane, and



(a) Carriageway **C**, maximum desired speed (b) Carriageway **D**, safe time headway

Fig. 4. Per-lane maximum desired speed and safe time headway distributions, as computed on the two M40 carriageways **C** and **D**, respectively. Velocity increases and headway decreases from the righmost to the leftmost lane.

hysteresis to avoid too frequent lane changes. Instead, we set two key parameters, *i.e.*, the maximum desired speed and the minimum safe time headway, on a per-vehicle basis. This is done in two steps: due to space limitations, we briefly sketch them below, and refer the reader to [7] for full details.

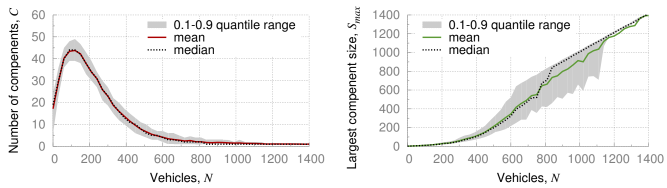
- We compute the distributions of the maximum speed and safe time headway of vehicles on each lane of the two carriageways, as recored at their respective entry points **51.2 C** and **57.0 D**. The former are derived by considering traffic in free-flow conditions only, *i.e.*, when drivers can keep their desired speed; the latter are computed using road traffic theory laws, from information on the vehicle density and average speed. Examples of these distributions are portrayed in Fig. 4.
- We tailor these distributions to each single vehicle, by truncating them at its recorded speed and time headway when entering the road segment. Drawing a random value from the tailored distributions allows associating to each vehicle a maximum desired speed and a minimum safe time headway that are consistent with (*i.e.*, respectively not lower nor higher than) the values recorded in the measurement data.

D. *A new dataset of bidirectional highway traffic*

We run simulations of road traffic on the target M40 segment in quasi-stationary conditions, by feeding our denoised measurement data to the microscopic-level mobility models parameterized on a per-vehicle basis. The resulting dataset describes five days of bidirectional highway traffic, and is the longest and most complete dataset of its kind to date.

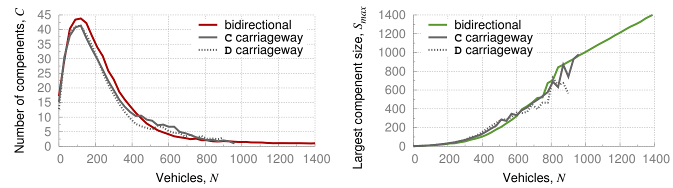
For the sake of reproducibility and verifiability, and to support reliability in the simulation of vehicular networks, we plan to make our simulator and the resulting road traffic datasets openly available to the research community[1].

[1]Simulator and data are at `http://www.it.uc3m.es/madrid-traces/`.



(a) C (b) S_{max}

Fig. 5. Dynamics of (a) C and (b) S_{max} versus N, on the M40 bidirectional motorway segment. Curves refer to the mean, median and 0.1-0.9 quantile range of C, as observed over the five days covered by our original dataset.



(a) C (b) S_{max}

Fig. 6. Dynamics of (a) C and (b) S_{max} versus N, on the M40 bidirectional motorway segment. Mean values are compared against those obtained when considering the unidirectional traffic in the two carriageways separately.

IV. Vehicular network connectivity: the three-phase law on bidirectional highways

We investigate the topological properties of vehicular networks supported by direct vehicle-to-vehicle communication, in the realistic highway environment granted by our original mobility dataset. Specifically, we focus on verifying if the three-phase law of connectivity proven for unidirectional highway traffic in [7] also holds also in presence of bidirectional traffic. We model the vehicular network as a graph, where vehicles map to nodes and communication links to edges. We then study how the number C of connected components (*i.e.*, graph subsets such that no edge exists between nodes in two different subsets) and the size of the maximum component S_{max} (*i.e.*, the number of vehicles in the largest of such subsets) vary with the number N of nodes in the graph.

Components map to clusters of vehicles that can communicate via multi-hop transmissions, and their number C is proportional to the fragmentation level of the vehicular network. The number of nodes is a proxy for the road traffic density. The three-phase connectivity law states the following.

I. For low values of N, C grows quasi-linearly and S_{max} slowly increases with the number of vehicles. That is, vehicles are isolated and create single-node components in sparse road traffic conditions.
II. Once a threshold in N is attained, C starts decreasing and S_{max} starts increasing faster with N. In other words, as the road traffic intensity grows beyond a critical value, additional vehicles join existing multi-hop clusters and improve the overall connectivity.
III. After a second N threshold, C equals one and S_{max} equals N, for any N. Beyond a specific road traffic volume, the vehicular network becomes fully connected.

The three-phase law remains valid also in the case of bidirectional traffic. Analyses with our original five-day dataset unveil that the three phases outlined above are exactly reproduced when simultaneously considering road traffic on the two opposite carriageways. This is shown in Fig. 5, which portrays the dynamics of C and S_{max} versus N. The result lets us speculate that the direction of traffic may have marginal impact on the vehicular network connectivity, as long as cars move in a quasi-unidimensional space such as that of highways.

In order to verify the conjecture above, we plot the mean behavior of C and S_{max} versus N in presence of bidirectional traffic, along with those recorded when considering each carriageway in isolation. Basically, we compare the average number of components and largest component size observed at a specific vehicular density, for bidirectional and unidirectional traffic respectively. Fig. 6 shows that the difference is minimal, and we ascribe it to a slightly more uniform spatial distribution of vehicles in the bidirectional case, where cars on opposite carriageways do not affect each other's speed. This leads to a connectivity that is worse for low N and better for high N.

Apart from such a minor diversity, the trends are consistent, which indicates that the three-phase law may hold in even more general unidimensional road traffic environments. For instance, we verified that this is the case under varying communication ranges: the results above refer to a range of 50 m, but the conclusions are identical under larger ranges.

V. Conclusions

In this short paper, we present original measurement data and use them to feed a properly calibrated simulation of microscopic vehicular mobility. The resulting dataset allows demonstrating that the three-phase law of vehicular connectivity in highway environments also holds in presence of bidirectional road traffic. For the sake of research reproducibility, the datasets employed in our work are made publicly available.

Acknowledgements

This research work has been performed in the framework of the H2020-ICT-2014-2 project 5G NORMA (Grant Agreement No. 671584), and of the People Programme (Marie Curie Actions) of the European Union's FP7/2007-2013 under REA grant agreement n.630211 ReFleX.

References

[1] US Department of Transportation, "Federal Motor Vehicle Safety Standards; V2V Communications, Notice of Proposed Rulemaking," National Highway Traffic Safety Administration, Dkt. No. NHTSA-2016-0126, 82 Fed. Reg. 3854, 2017.
[2] 5GPPP, "5G Automotive Vision," *White paper*, 2015.
[3] S. Uppoor, O. Trullols-Cruces, M. Fiore, J.Barcelo-Ordinas, "Generation and analysis of a large-scale urban vehicular mobility dataset," *IEEE Transactions on Mobile Computing*, 13(5), 2014.
[4] L. Codeca, R. Frank, T. Engel, "Luxembourg SUMO Traffic (LuST) Scenario: 24 hours of mobility for vehicular networking research," *IEEE VNC*, 2015.
[5] L. Bedogni, M. Gramaglia, A. Vesco, M. Fiore, J. Härri, F. Ferrero, "The Bologna Ringway dataset: improving road network conversion in SUMO and validating urban mobility via navigation services," *IEEE Transactions on Vehicular Technology*, 64(12), 2015.
[6] N. Akhtar, S. Coleri Ergen, O. Ozkasap, "Vehicle Mobility and Communication Channel Models for Realistic and Efficient Highway VANET Simulation," *IEEE Transactions on Vehicular Technology*, 64(1), 2015.
[7] M. Gramaglia, O. Trullols-Cruces, D. Naboulsi, M. Fiore, M. Calderon, "Mobility and connectivity in highway vehicular networks: a case study in Madrid," *Computer Communications*, 78(15), 2016.
[8] M. Treiber, A. Hennecke, D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E*, 62(2), 2000.
[9] M. Treiber, D. Helbing, "Realistische mikrosimulation von strassenverkehr mit einem einfachen modell," *ASIM*, 2002.